Protein labeling reactions in electrochemical microchannel flow: Numerical simulation and uncertainty propagation

Bert J. Debusschere[a] and Habib N. Najm[b]
Sandia National Labs, Livermore, California 94550

Alain Matta,[c] Omar M. Knio,[d] and Roger G. Ghanem[e]
The Johns Hopkins University, Baltimore, Maryland 21218

Olivier P. Le Maître[f]
Université d'Evry Val d'Essonne, Evry, France

(Received 30 October 2002; accepted 18 April 2003; published 30 June 2003)

This paper presents a model for two-dimensional electrochemical microchannel flow including the propagation of uncertainty from model parameters to the simulation results. For a detailed representation of electroosmotic and pressure-driven microchannel flow, the model considers the coupled momentum, species transport, and electrostatic field equations, including variable zeta potential. The chemistry model accounts for pH-dependent protein labeling reactions as well as detailed buffer electrochemistry in a mixed finite-rate/equilibrium formulation. Uncertainty from the model parameters and boundary conditions is propagated to the model predictions using a pseudo-spectral stochastic formulation with polynomial chaos (PC) representations for parameters and field quantities. Using a Galerkin approach, the governing equations are reformulated into equations for the coefficients in the PC expansion. The implementation of the physical model with the stochastic uncertainty propagation is applied to protein-labeling in a homogeneous buffer, as well as in two-dimensional electrochemical microchannel flow. The results for the two-dimensional channel show strong distortion of sample profiles due to ion movement and consequent buffer disturbances. The uncertainty in these results is dominated by the uncertainty in the applied voltage across the channel. [DOI: 10.1063/1.1582857]

I. INTRODUCTION

Microchannel flows, involving electroosmotic flow of charged components in an electrolyte buffer, are generally characterized by strong coupling between multiple physical and chemical processes.[1] Numerical simulations for detailed studies of phenomena such as analyte dispersion therefore require accurate models for the fluid flow, species transport, chemical reactions, buffer equilibrium, protein ampholytic behavior, electrostatic field strength, wall layer, and many other processes.[2] Most of these processes are well understood and adequate models are generally available. Many simulations of microchannel flow can be found in the literature, with varying detail in the resolution of the ongoing physical processes.[3–13] However, simulations that take into account the full range of coupled processes in microchannel flows are hard to find.

Further, the mathematical models for these physical processes generally require knowledge of several parameters such as species mobilities, viscosity, electrolyte dissociation constants, reaction rate parameters, and other physical and environmental parameters. These parameters are typically not known exactly due to experimental measurement uncertainties and/or inherent variability. Consequently, computational predictions will have some uncertainty, associated with the uncertainties in the input parameters. In order to make valid comparisons between experimental and computational data, or to assess the reliability of computational predictions, a careful analysis of the uncertainty in the simulation results is required.

In the current work, a detailed model is developed of both electroosmotic and pressure-driven flow in a microchannel filled with an electrolyte buffer and model protein analyte samples. The construction considers the fully coupled momentum, species transport, and electrostatic field equations, including a model for the dependence of the zeta potential on pH and buffer molarity. A mixed finite-rate, partial-equilibrium formulation is applied for the chemical reactions. In particular, ''fast'' electrolyte reactions are described by associated equilibrium constraints, while the remaining ''slow'' protein labeling reactions are modeled with finite-rate kinetics.

To quantify the uncertainty in the model predictions, due to uncertainty in the input parameters, a stochastic uncertainty propagation method[14] is applied. This method introduces a new stochastic dimension for each uncertain parameter, and uses polynomial chaos (PC) expansions[15] to describe the dependence of model parameters and flow quan-

[a]Electronic mail: bjdebus@ca.sandia.gov
[b]Electronic mail: hnnajm@ca.sandia.gov
[c]Electronic mail: alan@mars.ce.jhu.edu
[d]Electronic mail: knio@jhu.edu
[e]Electronic mail: ghanem@jhu.edu
[f]Electronic mail: olm@iup.univ-evry.fr

tities on these dimensions. After introducing these PC representations into the governing equations, a Galerkin approach is used to determine evolution equations for the spectral mode strengths in the expansion. The resulting system is more complex than the corresponding deterministic model, requiring more computational effort. However, it is potentially more efficient than Monte Carlo (MC) simulations.[16] Moreover, the pseudo-spectral PC approach readily provides sensitivity information and the contribution to the total uncertainty by each of the model parameters.

First we outline the formulation of the governing equations that constitute the deterministic system model. Next, we implement the stochastic uncertainty quantification method to reformulate these equations into evolution equations for the spectral mode strengths. We then proceed to the description of the numerical construction used to integrate the resulting set of equations, highlighting particular developments necessary for handling the coupled evolution of momentum, species, and the electrostatic field. The methodology is then applied to model protein labeling reactions in homogeneous systems as well as two-dimensional microchannel flows. The results illustrate the convergence of the construction as well as the propagation/growth of uncertainty in the simulations. The detailed physical model gives insight into important microfluidic sample dispersion mechanisms.

II. PHYSICAL MODEL FORMULATION

A. Momentum

The continuity and momentum equations for a two-dimensional flow field in the (x,y) plane, with uniform density and viscosity are given by[17]

$$\nabla\cdot\mathbf{u}=0, \tag{1}$$

$$\frac{\partial\mathbf{u}}{\partial t}+\mathbf{u}\cdot\nabla\mathbf{u}=-\nabla p+\nu\nabla^2\mathbf{u}, \tag{2}$$

where $\mathbf{u}$ is the velocity, p is the pressure normalized by density, and ν is the kinematic viscosity.

The microchannel flows in this study are electroosmotically driven with an applied electrostatic field in the x direction. Assuming a double layer that is thin with respect to the channel size, the effect of wall electrostatic forces can be represented in terms of a wall slip velocity u_w, using the Helmholtz–Smoluchowski relationship[17]

$$u_w=\frac{\epsilon\zeta}{\mu}\nabla_t\phi_w, \tag{3}$$

where ϵ is the permittivity of the fluid, ζ is the zeta potential, ϕ_w is the electrostatic field potential at the wall, and μ is the dynamic viscosity. Since both the electrostatic field and the ζ potential depend on the fluid composition, Eq. (3) represents a major coupling between the flow velocity and the species transport.

The ζ potential is a function of the wall material and fluid characteristics.[7,18] In this work, a relationship for ζ as a function of the local pH and buffer molarity was obtained from empirical data for the zeta potential of a fused silica capillary in an aqueous solution of KCl, as shown in Fig. 1.[19]

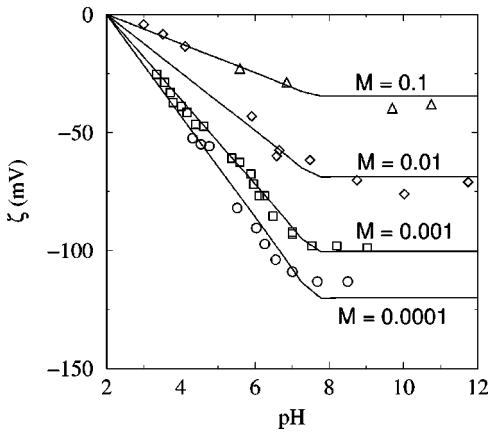

FIG. 1. Empirical data and curve fit for the ζ potential of a fused silica capillary vs pH in an aqueous solution of KCl at various molarities. Adapted with permission from Ref. 19, Copyright 1992 American Chemical Society.

These data were curve-fitted into the following relationship:

$$\zeta(\mathrm{pH},\mathcal{M}) =\{-(\mathrm{pH}-2)+(\tfrac{1}{2}+\tfrac{1}{2}\tanh(5(\mathrm{pH}-7.5)))(\mathrm{pH}-7.6)\} \times(-2.7\ln(\mathcal{M}+2.3\times10^{-4})), \tag{4}$$

where $\mathcal{M}$ is the molarity of the KCl solution. The quantitative accuracy of this curve-fit is obviously limited to systems similar to the one considered in Ref. 19. However, Eq. (4) qualitatively gives the correct behavior of $\zeta(\mathrm{pH},\mathcal{M})$ for various other systems.[7,18]

B. Species concentrations

A variety of species are considered in this work, ranging from model proteins and dyes in samples, to the ions of aqueous buffer solutions. The transport of these species is governed by[17]

$$\frac{\partial c_i}{\partial t}+\nabla\cdot[c_i(\mathbf{u}+\mathbf{u}_i^e)]=\nabla\cdot(D_i\nabla c_i)+\hat{w}_i, \tag{5}$$

where c_i is the concentration of species i, and D_i is the corresponding diffusivity. The electromigration velocity $\mathbf{u}_i^e$ accounts for the electrophoretic movement of electrically charged species relative to the bulk flow. This velocity is given by[17]

$$\mathbf{u}_i^e=-\beta_i z_i F\nabla\phi, \tag{6}$$

where β_i is the electrophoretic mobility for species i, z_i is the charge number, F is the Faraday constant (9.648×10^4 C/mol), and ϕ is the electrostatic field potential. The term $\hat{w}_i$ is a source term from the chemical and electrochemical reactions in which species i is involved. Note that for each species, the diffusivity D_i and the mobility β_i are coupled through the Nernst–Einstein equation[17]

$$D_i=RT\beta_i, \tag{7}$$

where R is the universal gas constant and T the temperature.

The integration of Eq. (5) is performed differently depending on the chemical time scales involved. In general,

electrolyte association and dissociation reaction rates are several orders of magnitude faster than electrophoretic phenomena[1] and typical sample-processing reactions. Thus, direct integration of fast reactions would impose severe time step restrictions. In order to avoid these difficulties, an equilibrium approach for the electrolyte reactions is implemented. For example, consider a weak acid HA, which dissociates according to

$$\mathrm{HA} \overset{K_\mathrm{A}}{\leftrightarrow} \mathrm{H}^+ + \mathrm{A}^-, \tag{8}$$

where

$$K_\mathrm{A} \equiv \frac{[\mathrm{H}^+][\mathrm{A}^-]}{[\mathrm{HA}]} \tag{9}$$

is the corresponding dissociation constant. Instead of integrating Eq. (5) for the concentrations of species HA and A^- individually, consider the combined concentration of both of these quantities $\theta_a = [\mathrm{HA}] + [\mathrm{A}^-]$. The source terms for $[\mathrm{HA}]$ and $[\mathrm{A}^-]$ from the electrolyte reaction (8) cancel out in the θ_a transport equation, which is the sum of the transport equations for the two individual quantities,

$$\frac{\partial \theta_a}{\partial t} + \nabla\cdot[c_{\mathrm{HA}}(\mathbf{u}+\mathbf{u}^e_{\mathrm{HA}}) + c_{\mathrm{A}^-}(\mathbf{u}+\mathbf{u}^e_{\mathrm{A}^-})] = \nabla\cdot[D_{\mathrm{HA}}\nabla c_{\mathrm{HA}} + D_{\mathrm{A}^-}\nabla c_{\mathrm{A}^-}]. \tag{10}$$

Therefore, barring any other chemical reactions involving these species, θ_a is a conserved quantity and can be integrated with Eq. (10) without a chemical source term.[2,4,10] Note that if the chemical source terms for HA or A^- in Eq. (5) do include participation by reactions other than the HA buffer chemistry, e.g., by (typically slow) sample chemistry, then the utilization of θ_a is still advantageous in that it eliminates the fast electrolyte reactions, but in this case θ_a is no longer a conserved scalar. In either case, one arrives at a governing equation for θ_a, which does not include the fast reaction terms. Once θ_a is known, the concentrations of the individual components of the weak acid are obtained from

$$[\mathrm{HA}] = \frac{[\mathrm{H}^+]}{[\mathrm{H}^+]+K_\mathrm{A}}\theta_a \equiv \alpha_{\mathrm{HA}} \times \theta_a, \tag{11}$$

$$[\mathrm{A}^-] = \frac{K_\mathrm{A}}{[\mathrm{H}^+]+K_\mathrm{A}}\theta_a \equiv \alpha_{\mathrm{A}^-} \times \theta_a. \tag{12}$$

Note that this construction is equally useful for buffers with multiple dissociation states, where θ_a is the sum of concentrations of the weak acid and all of its dissociated states. Since the mobilities and diffusivities are generally different for the species that make up θ_a, the convection and diffusion terms in the transport equation for θ_a are calculated as the sum of the convection and diffusion for each species in θ_a. A similar approach holds for weak bases.

For strong acids and bases, which are fully dissociated in the solution, or for other species that do not take part in electrolyte dissociation and association reactions, Eq. (5) can be integrated directly. The model proteins and fluorescent dyes in this work are assumed to have a fixed charge, so their concentrations are integrated using Eq. (5), with an appropriate finite-rate chemical source term. However, a complete ampholyte description for proteins can readily be formulated with a similar framework as is used for the weak acids and bases.[2,20–22] In the simulations in this work, proteins are assumed to take part in a finite rate, irreversible labeling reaction of the form

$$\mathrm{U} + \mathrm{D} \overset{k_\mathrm{L}}{\rightarrow} \mathrm{L} \tag{13}$$

with a pH-dependent reaction rate $k_\mathrm{L} = k_\mathrm{L}(\mathrm{pH})$. In Eq. (13), U is the unlabeled protein, D the fluorescent dye, and L the labeled protein.

Since a thin double layer is assumed, the system is also assumed to satisfy the electroneutrality condition

$$\sum_i z_i c_i = 0 \tag{14}$$

everywhere in the domain. The concentrations of H^+ and OH^- are obtained from this electroneutrality condition and the water dissociation constant

$$[\mathrm{H}^+][\mathrm{OH}^-] = K_w. \tag{15}$$

Note that the composition, and therefore also the total charge, of weak acids and bases in the system depends on the H^+ concentration [see Eqs. (11) and (12) above]. The substitution of Eqs. (12) and (15) into the electroneutrality condition (14), in order to account for the dependence of $[\mathrm{A}^-]$ and $[\mathrm{OH}^-]$ on $[\mathrm{H}^+]$, introduces nonlinear terms in this equation. For buffers with multiple dissociation states, even more nonlinear terms are introduced. Therefore, an iterative solution of the electroneutrality condition for $[\mathrm{H}^+]$ is usually required.

C. Electrostatic field strength

Allowing for concentration field gradients, the electrostatic field potential, ϕ, is obtained from the current continuity constraint,[17]

$$\nabla\cdot(\sigma\nabla\phi) = -F\sum_i z_i \nabla\cdot(D_i \nabla c_i). \tag{16}$$

This equation is coupled to the species concentrations through the right-hand side (diffusion of charge) and the electrical conductivity σ of the solution

$$\sigma = F^2 \sum_i z_i^2 \beta_i c_i. \tag{17}$$

The electrostatic field strength is then obtained as $\mathbf{E} = -\nabla\phi$.

This completes the description of the key elements of the deterministic model formulation. We next outline the stochastic construction for uncertainty quantification, and its implementation in this model.

III. STOCHASTIC FORMULATION

To propagate uncertainty from the input parameters of the physical model, to the results of model simulations, we use a spectral stochastic formulation based on polynomial chaos (PC) expansions.[14,15] For each uncertain parameter,

this technique introduces a new stochastic dimension ξ, where ξ is a random variable with a standard normal Gaussian probability density function (PDF). The parameter is then modeled as a random variable whose dependence on ξ is described using a spectral PC expansion. For example, the species diffusivity D is written as

$$D(\xi)=\sum_{k=0}^{P} D_k\Psi_k(\xi), \tag{18}$$

where the Ψ_k are the PC basis functions and the deterministic coefficients D_k are the spectral mode strengths. If the model has only one uncertain parameter, then the basis functions are the one-dimensional Hermite polynomials:

$$\Psi_0=1,\ \ \Psi_1=\xi,\ \ \Psi_2=\xi^2-1,\ \ \Psi_3=\xi^3-3\xi,\ \ldots. \tag{19}$$

In this case, P corresponds to the highest order polynomial used in the expansion. For a model with N uncertain parameters, an N-dimensional stochastic space is considered, and the Ψ_k are N-dimensional Hermite polynomials up to a specified order p in the Gaussian variables $\theta=\{\xi_1,\xi_2,\ldots,\xi_N\}$,[14] such that

$$D(\theta)=\sum_{k=0}^{P} D_k\Psi_k(\theta). \tag{20}$$

Note that in this case, the Gaussian variables $\xi_1,\ldots,\xi_N$ are assumed to be uncorrelated. The number of terms $(P+1)$ in the expansion corresponds to the number of polynomials with order less or equal to p. Higher order polynomial chaos basis functions increase the accuracy of the spectral representation, but add a lot more terms for problems with many stochastic dimensions.

The solution field variables, such as velocities, concentrations, and the electrostatic field potential are expanded similarly:

$$\mathbf{u}(\mathbf{x},t;\theta)=\sum_{k=0}^{P} \mathbf{u}_k(\mathbf{x},t)\Psi_k(\theta), \tag{21}$$

$$c(\mathbf{x},t;\theta)=\sum_{k=0}^{P} c_k(\mathbf{x},t)\Psi_k(\theta), \tag{22}$$

$$\phi(\mathbf{x},t;\theta)=\sum_{k=0}^{P} \phi_k(\mathbf{x},t)\Psi_k(\theta). \tag{23}$$

Using the orthogonality of the basis functions, the coefficients D_k, $\mathbf{u}_k(\mathbf{x},t)$, $c_k(\mathbf{x},t)$, and $\phi_k(\mathbf{x},t)$ can be calculated by a projection operation onto the PC basis. Thus, for D_k

$$D_k=\frac{\langle\Psi_k D\rangle}{\langle\Psi_k^2\rangle}, \tag{24}$$

where the expectation is defined as

$$\langle f\rangle=\frac{1}{(2\pi)^{N/2}}\int_{R^N} f(\xi_1,\xi_2,\ldots,\xi_N)\times\exp\left[-\frac{\Sigma_{i=1}^{N}\xi_i^2}{2}\right]d\xi_1 d\xi_2\cdots d\xi_N. \tag{25}$$

Note that the expectations $\langle\Psi_k\rangle=0$ for $k>0$. Therefore the zeroth-order spectral mode for each field quantity represents the mean field, whereas the higher order modes represent the variation, or uncertainty, around this mean. This is reflected in the calculation of the standard deviation σ of a quantity, as shown below for the concentration c:

$$\begin{aligned}\sigma_c^2&=\langle(c-\langle c\rangle)^2\rangle\\&=\left\langle\left(\sum_{j=1}^{P} c_j\Psi_j\right)\left(\sum_{k=1}^{P} c_k\Psi_k\right)\right\rangle\\&=\sum_{j=1}^{P}\sum_{k=1}^{P} c_j c_k\langle\Psi_j\Psi_k\rangle\\&=\sum_{k=1}^{P} c_k^2\langle\Psi_k^2\rangle.\end{aligned} \tag{26}$$

To further interpret this, consider the example of a second-order PC expansion for the concentration c in the case of two uncertain parameters:

$$c=c_0+c_1\xi_1+c_2\xi_2+c_3(\xi_1^2-1)+c_4\xi_1\xi_2+c_5(\xi_2^2-1). \tag{27}$$

In Eq. (27), the ξ_1 dimension corresponds to the first uncertain parameter in the system, and ξ_2 corresponds to the second uncertain parameter. Substituting this expression into Eq. (26) and evaluating the $\langle\Psi_k^2\rangle$, we obtain the following equation for the variance of c in this example:

$$\sigma_c^2=c_1^2+c_2^2+2c_3^2+c_4^2+2c_5^2. \tag{28}$$

Grouping the terms that correspond to the same stochastic dimensions, it is possible to identify the contributions of individual parameters to this overall standard deviation:

$$\sigma_c^2=\sigma_1^2+\sigma_{12}^2+\sigma_2^2, \tag{29}$$

$$\sigma_1=\sqrt{c_1^2+2c_3^2}, \tag{30}$$

$$\sigma_{12}=\sqrt{c_4^2}, \tag{31}$$

$$\sigma_2=\sqrt{c_2^2+2c_5^2}. \tag{32}$$

In these expressions, σ_1 and σ_2 represent the individual contributions of the first and second uncertain parameters, respectively, to the overall uncertainty in c. The term σ_{12} represents a coupled term involving the combined effect of the two parameters. This feature of the PC methodology is very powerful, as it allows identification of the major contributions to the uncertainty in the simulation output and model results.

After representing all model parameters and solution fields with PC expansions, those expansions are substituted into the transport equations for the deterministic quantities. Evolution equations for the unknown coefficients in the PC expansions are then obtained by a Galerkin approach. For example, consider the momentum equation. Substituting the appropriate PC expansions for velocities, pressure, and viscosity in Eq. (2), multiplying by Ψ_k, and taking the expectation gives[23,24]

$$\frac{\partial \mathbf{u}_k}{\partial t}+\sum_{i=0}^{P}\sum_{j=0}^{P} C_{ijk}(\mathbf{u}_i\cdot\nabla)\mathbf{u}_j = -\nabla p_k+\sum_{i=0}^{P}\sum_{j=0}^{P} C_{ijk}\nu_i\nabla^2\mathbf{u}_j \tag{33}$$

with $C_{ijk}\equiv\langle\Psi_i\Psi_j\Psi_k\rangle/\langle\Psi_k\Psi_k\rangle$. Similarly, the equations for the modes $c_{m,k}$ of the species concentration (with m the species index) become

$$\frac{\partial c_{m,k}}{\partial t}+\sum_{i=0}^{P}\sum_{j=0}^{P} C_{ijk}\nabla\cdot(c_{m,i}(\mathbf{u}_j+\mathbf{u}^e_{m,j})) = \sum_{i=0}^{P}\sum_{j=0}^{P} C_{ijk}\nabla\cdot(D_{m,i}\nabla c_{m,j})+\hat{w}_{m,k}, \tag{34}$$

where

$$\mathbf{u}^e_{m,j}=\frac{\langle\Psi_j\mathbf{u}^e\rangle}{\langle\Psi_j^2\rangle}=\sum_{k=0}^{P}\sum_{i=0}^{P} C_{kij}\beta_k zF\nabla\phi_i, \tag{35}$$

$$\hat{w}_{m,k}=\frac{\langle\Psi_k\hat{w}\rangle}{\langle\Psi_k^2\rangle}. \tag{36}$$

Equations (35) and (36) represent the pseudo-spectral projection of the electrophoretic velocities and the stochastic chemical source terms onto the ψ_k basis functions (see also Sec. IV F). Finally, the electrostatic field equation (16) becomes

$$\sum_{i=0}^{P}\sum_{j=0}^{P} C_{ijk}\nabla\cdot(\sigma_i\nabla\phi_j) = -F\sum_m z_m\sum_{i=0}^{P}\sum_{j=0}^{P} C_{ijk}\nabla\cdot(D_{m,i}\nabla c_{m,j}). \tag{37}$$

The modes σ_i of the electrical conductivity are obtained from

$$\sigma_i=F^2\sum_m z_m^2\sum_{j=0}^{P}\sum_{k=0}^{P} C_{jki}\beta_{m,j}c_{m,k}. \tag{38}$$

Equations (33), (34), and (37) each represent a set of $P+1$ coupled equations to be solved for the mode strengths u_k, $c_{m,k}$, and ϕ_k, $k=0,\ldots,P$. With M species, the total number of equations to be solved is $(M+2)(P+1)$.

IV. IMPLEMENTATION

A. Data structure

As described in Sec. II B, species concentrations are integrated differently, based on whether or not they take part in equilibrium reactions. For instance, for components of weak acids or bases, which typically serve as buffers, only the combined concentration of all components is integrated directly. The total charge associated with the buffer components is required for the enforcement of the electroneutrality equation (14). For a given buffer, this total charge can be obtained from the total buffer concentration θ and $[H^+]$ through buffer-specific equations such as Eqs. (11) and (12).

To make the treatment of weak acids or bases as general as possible, separate objects are used in the current code to represent these components. Each object contains all the species properties for the weak acid or base it represents, as well as the dissociation constants for the electrolyte reactions between its species. Specific functions are also associated with each object to return the total charge or other information about the weak acid or base, given its total concentration and $[H^+]$. This way, different buffers can be included in the simulations by simply including different objects, without the need for specific code modifications.

B. Spatial discretization

The computational domain is discretized using a Cartesian mesh with uniform cell size Δx and Δy in the x and y direction, respectively. Vector fields, such as the velocity and the electrostatic field strength, are defined on the cell faces. Scalar fields, such as pressure and species concentrations, are defined at the cell centers. Spatial derivatives are discretized with second-order central differences.

C. Electroneutrality

As explained in Sec. II B, the individual concentrations of the buffer ions and $[H^+]$ are obtained from the electroneutrality condition (14). This results in a set of nonlinear algebraic relations between $P+1$ stochastic modes. This coupled nonlinear system of equations is iteratively solved at each point in the domain, using a Newton solver from the NITSOL package.[25] The solver uses an inexact Newton method with backtracking. Using the solution from the previous time step as initial guess, the convergence is generally very fast.

D. Electrostatic field strength

To obtain the electrostatic field potential ϕ, the set of $P+1$ equations (37) needs to be solved over the domain. Since these equations are coupled, an iterative solution method was developed, consisting of Gauss–Seidel iterations over the spatial dimensions in combination with SOR iterations over the stochastic dimensions. To accelerate the convergence speed, spatial coarsening with a multigrid approach is applied. The electrostatic field strength is computed in turn as the gradient of the electrostatic potential.

E. Time integration

The time integration algorithm in this work is based on a previously developed stochastic projection method for the momentum equations in low-Mach-number flow.[24] This momentum solver uses a time splitting approach in which the convection and diffusion terms are integrated in a first fractional step, and the continuity constraints are then enforced in a pressure projection step.[26] Since the continuity constraints [Eq. (1)] are decoupled in the stochastic dimension, this leads to a set of $P+1$ decoupled Poisson problems.

In the current work, this method is expanded to the integration of the coupled momentum and species transport equations, in combination with the electrostatic field solu-

tion. For brevity, the equations for the stochastic mode k of the species concentrations and the velocity can be written as

$$\frac{\partial c_k}{\partial t} = -\mathcal{C}sp_k + \mathcal{D}sp_k + \mathcal{S}sp_k, \tag{39}$$

$$\frac{\partial \mathbf{u}_k}{\partial t} = -\mathcal{C}m_k + \mathcal{D}m_k - \nabla p_k \tag{40}$$

where $\mathcal{C}sp_k$, $\mathcal{D}sp_k$, and $\mathcal{S}sp_k$ represent the convection, diffusion, and chemical source terms in the species equation (34). Similarly, $\mathcal{C}m_k$ and $\mathcal{D}m_k$ represent the convection and diffusion terms in the momentum equation (33). Using the projection scheme for momentum, in combination with a Runge–Kutta (RK) time integration scheme, Eqs. (39) and (40) are discretized between t^n and the RK stage time level $t^{(s)} = t^n + \Delta t^{(s)}$ as

$$\frac{c_k^{(s)} - c_k^n}{\Delta t^{(s)}} = -\mathcal{C}sp_k^{(s-1)} + \mathcal{D}sp_k^{(s-1)} + \mathcal{S}sp_k^{(s-1)} \equiv \mathcal{F}sp_k^{(s-1)}, \tag{41}$$

$$\frac{\mathbf{u}_k^{(s),*} - \mathbf{u}_k^n}{\Delta t^{(s)}} = -\mathcal{C}m_k^{(s-1)} + \mathcal{D}m_k^{(s-1)} \equiv \mathcal{F}m_k^{(s-1)}, \tag{42}$$

$$\frac{\mathbf{u}_k^{(s)} - \mathbf{u}_k^{(s),*}}{\Delta t^{(s)}} = -\nabla p_k^{(s)}, \tag{43}$$

where $\mathcal{F}sp_k$ and $\mathcal{F}m_k$ represent the full right-hand sides in the corresponding time integration steps. Equation (43) is the pressure correction step, which requires the pressure to be solved for first. The equation for pressure is obtained by substituting Eq. (43) into the stochastic form of the continuity equation for $\mathbf{u}^{(s)}$,

$$\nabla \cdot \mathbf{u}_k^{(s)} = 0 \tag{44}$$

resulting in the following set of Poisson equations:

$$\nabla^2 p_k^{(s)} = -\frac{1}{\Delta t^{(s)}} \nabla \cdot \mathbf{u}_k^{(s),*} \qquad k = 0, \ldots, P. \tag{45}$$

As discussed in Ref. 24, these $P+1$ Poisson equations are decoupled; therefore, each can be solved individually using existing Poisson solvers for deterministic flow problems. In the current work, the same fast Fourier transform based flow solver is used as in Ref. 24.

The time integration of Eqs. (41) and (42) is performed using the four-stage, fourth order Runge–Kutta scheme (RK4),[27] which was selected because of its good stability for convection dominated problems. Keeping in mind the coupling between the equations, the computations during the subsequent stages of the RK4 integration over a time step Δt from time t^n to $t^{n+1} = t^n + \Delta t$ can be represented with the following pseudo-code. The superscripts (s) denote the Runge–Kutta stage number. For clarity, the subscripts for the mode strength k have been dropped.

Stage $s = 1$; $t = t^n$.

Calculate the right hand sides in Eqs. (41) and (42) using the species concentrations, velocities and electrostatic field strength at time $t = t^n$:

- $\mathcal{F}sp^{(1)} = \mathcal{F}sp(c(t^n), \mathbf{u}(t^n), \mathbf{E}(t^n))$,
- $\mathcal{F}m^{(1)} = \mathcal{F}m(\mathbf{u}(t^n), u_w(t^n))$ where u_w is the electroosmotic wall velocity.

Stage $s = 2,3,4$; $t = t^n + \Delta t^{(s)}$.

Update species concentrations to the current time level:

- $c^{(s)} = c(t^n) + \Delta t^{(s)} \mathcal{F}sp^{(s-1)}$ for all directly integrated species.
- Solve electroneutrality constraint to obtain $[\mathrm{H}^+]^{(s)}$.
- Update concentrations of weak acids and/or bases.

Update electrostatic field strength and velocity boundary conditions using the updated concentrations:

- $\mathbf{E}^{(s)} = \mathbf{E}(c^{(s)})$.
- $u_w^{(s)} = u_w(c^{(s)}, \mathbf{E}^{(s)})$.

Update velocities to the current time level:

- Update the velocities to their intermediate (*) values at the current time level: $\mathbf{u}^{(s),*} = \mathbf{u}(t^n) + \Delta t^{(s)} \mathcal{F}m^{(s-1)}$.
- Apply the boundary conditions $u_w^{(s)}$ to the $\mathbf{u}^{(s),*}$ velocity field.
- Solve for pressure at this time level using Eq. (45): $p^{(s)} = p(\mathbf{u}^{(s),*})$.
- Apply the pressure correction to $\mathbf{u}^{(s),*}$ to obtain $\mathbf{u}^{(s)}$: $\mathbf{u}^{(s)} = \mathbf{u}^{(s),*} - \Delta t^{(s)} \nabla p^{(s)}$.

Calculate the new right hand sides in Eqs. (41) and (42) using the updated species concentrations, velocities, and electrostatic field strength:

- $\mathcal{F}sp^{(s)} = \mathcal{F}sp(c^{(s)}, \mathbf{u}^{(s)}, \mathbf{E}^{(s)})$,
- $\mathcal{F}m^{(s)} = \mathcal{F}m(\mathbf{u}^{(s)}, u_w^{(s)})$.

Final update to time $t^{n+1} = t^n + \Delta t$.

Update species concentrations to t^{n+1}:

- $c(t^{n+1}) = c(t^n) + \Delta t(\frac{1}{6}\mathcal{F}sp^{(1)} + \frac{2}{6}\mathcal{F}sp^{(2)} + \frac{2}{6}\mathcal{F}sp^{(3)} + \frac{1}{6}\mathcal{F}sp^{(4)})$ for all directly integrated species.
- Solve electroneutrality constraint to obtain $[\mathrm{H}^+]$ at t^{n+1}.
- Update concentrations of weak acids and/or bases.

Update electrostatic field strength and velocity boundary conditions using the updated concentrations:

- $\mathbf{E}(t^{n+1}) = \mathbf{E}(c(t^{n+1}))$,
- $u_w(t^{n+1}) = u_w(c(t^{n+1}), \mathbf{E}(t^{n+1}))$.

Update velocities to t^{n+1}:

- Update the velocities to the intermediate (*) values at t^{n+1}: $\mathbf{u}^*(t^{n+1}) = \mathbf{u}(t^n) + \Delta t(\frac{1}{6}\mathcal{F}m^{(1)} + \frac{2}{6}\mathcal{F}m^{(2)} + \frac{2}{6}\mathcal{F}m^{(3)} + \frac{1}{6}\mathcal{F}m^{(4)})$.
- Apply the boundary conditions $u_w(t^{n+1})$ to the $\mathbf{u}^*(t^{n+1})$ velocity field.

- Solve for pressure at t^{n+1} using Eq. (45): $p(t^{n+1}) = p(\mathbf{u}^*(t^{n+1}))$.
- Apply the pressure correction to $\mathbf{u}^*(t^{n+1})$ to obtain $\mathbf{u}(t^{n+1})$: $\mathbf{u}(t^{n+1}) = \mathbf{u}^*(t^{n+1}) - \Delta t \nabla p(t^{n+1})$.

In the above integration scheme, the respective time steps $\Delta t^{(s)}$ of the Runge–Kutta stages $s=2$, 3, and 4 are given by $\frac{1}{2}\Delta t$, $\frac{1}{2}\Delta t$, and Δt.

F. Tools for stochastic operations: Uncertainty quantification toolkit

As explained in Sec. III, the governing equations for the spectral mode strengths of the field variables are obtained by substituting the PC expansions for those field variables in their original, deterministic governing equations. Instead of explicitly writing out these equations for the spectral mode strengths, it is also possible to retain the governing equations in their original form, but take into account during the implementation that all arithmetic needs to be performed on stochastic instead of deterministic variables.

To facilitate this approach, we developed an uncertainty quantification (UQ) toolkit which contains subroutines to perform most of the common operations on stochastic variables that are represented by PC expansions. Using this toolkit, many algorithms that were originally designed for deterministic problems can easily be converted for stochastic computations by merely replacing mathematical operators with calls to their stochastic equivalent. The details of some of these operations are explained in the following.

Aside from additions, one of the most common operations is the multiplication of two stochastic variables. Consider two stochastic variables, u and v, with the following PC representations:

$$u = \sum_{i=0}^{P} u_i \Psi_i, \tag{46}$$

$$v = \sum_{j=0}^{P} v_j \Psi_j. \tag{47}$$

We need to find the modes w_k in the PC representation of $w = uv$:

$$w = \sum_{k=0}^{P} w_k \Psi_k. \tag{48}$$

As mentioned before, these coefficients are obtained by using the orthogonality property of the PC basis functions:

$$w_k = \sum_{i=0}^{P} \sum_{j=0}^{P} C_{ijk} u_i v_j, \quad k = 0, \ldots, P \tag{49}$$

with

$$C_{ijk} \equiv \frac{\langle \Psi_i \Psi_j \Psi_k \rangle}{\langle \Psi_k^2 \rangle}. \tag{50}$$

Since the tensor C_{ijk} is a function of the PC basis functions only, it only needs to be calculated once during a preprocessing step and can then be stored for use throughout the computations. The implementation of Eq. (49) also takes advantage of the fact that this tensor is sparse, reducing the amount of storage and CPU time needed.

A similar procedure could also be used to determine the PC expansion for the product of three stochastic variables $g = uvw$. This would give the spectral coefficients g_l as

$$g_l = \sum_{i=0}^{P} \sum_{j=0}^{P} \sum_{k=0}^{P} D_{ijkl} u_i v_j w_k, \quad l = 0, \ldots, P, \tag{51}$$

where

$$D_{ijkl} \equiv \frac{\langle \Psi_i \Psi_j \Psi_k \Psi_l \rangle}{\langle \Psi_l^2 \rangle}. \tag{52}$$

Instead of this pure spectral approach, however, a pseudo-spectral approach is used in this work to calculate products such as $g = uvw$ by repeated use of the regular product function. First the product uv is calculated with Eq. (49), and the result of this multiplication is multiplied in the same way with w to give the PC expansion for g. The advantage of this pseudo-spectral approach is that it does not require the evaluation and storage of the fourth-rank tensor D_{ijkl}, is more efficient, and is easy to generalize to products of any number of variables. Some aliasing errors are introduced though in this approach, but they were found to be negligible as long as the order of the PC expansions is chosen sufficiently high.

Another frequent operation is the calculation of the inverse of a stochastic quantity. To explain how this operation is implemented, consider again three stochastic variables, u, v, and w, with their respective PC expansions given by Eqs. (46)–(48). If we wish to calculate $u = w/v$, then this implies $w = uv$, which is given by Eq. (49). This equation, assuming the modes w_k and v_j are known, represents a system of $P+1$ linear equations in the unknown modes u_i. Since it is a sparse system of equations, it is solved efficiently in this work with a GMRES iterative solver, taken from the SLATEC library.[28]

More challenging is the evaluation of nonpolynomial functions of stochastic variables such as the exponential, which will show up in the calculation of the protein labeling reaction rate with Eq. (56), or the logarithm in the calculation of pH. Currently, these operations are performed by expanding them in Taylor series around the mean of the argument. For example, the exponential of a stochastic quantity u, with a PC expansion given by Eq. (46), is computed as

$$e^u = e^{u_0} \left(1 + \sum_{n=1}^{N} \frac{d^n}{n!} \right), \tag{53}$$

where

$$d = u - u_0 = \sum_{i=1}^{P} u_i \Psi_i \tag{54}$$

is the stochastic part of u. The powers d^n are again calculated in a pseudo-spectral way with the product formula (49), as $d^n = d\, d^{n-1}$, with d^{n-1} known from the previous term in the Taylor series. The number of terms N in this truncated series is chosen adaptively to satisfy a given tolerance level.

The Taylor series approach works reasonably well as long as the uncertainties in the field variables are moderate and the probability density functions (PDFs) of those variables are not too skewed. For highly skewed PDFs, however, high order PC expansions are required to capture this stochastic information, and the evaluation of high power terms d^n in the Taylor series can become inaccurate. A new approach, based on integrations, is currently under development to alleviate this problem.

All the operations described above, among many others, have been implemented in the UQ toolkit library. The UQ toolkit greatly facilitates the development of stochastic solvers from scratch, as well as the conversion of existing deterministic routines into stochastic ones. A more general and effective approach would be to develop the capability to automatically convert existing deterministic programs to stochastic arithmetic. As an extension of the UQ toolkit methodology, this could conceptually be achieved by creating a new data type for variables represented with PC expansions, and then overloading operators to perform the proper stochastic operations on these variables where needed. Such an approach is outside the scope of this paper but will be explored elsewhere.

V. NUMERICAL RESULTS

A. Test system

This section presents some results of test problems illustrating the spatial and temporal convergence properties of the developed code. Figure 2 shows the geometry considered for these test problems, consisting of a rectangular microchannel in which a protein U and dye D react to form a labeled protein L. An external electrostatic potential is applied across the system to generate an electroosmotic flow in the x direction. The unlabeled protein U has a charge of $+1$ versus a charge of -1 for the dye D, so electrophoresis will move U forward and D backward, relative to the bulk flow. For all cases simulated in this work, an aqueous potassium phosphate (KH_2PO_4) buffer solution is considered. Therefore, the species in the solution are the proteins U and L, the dye D, the electrolytes H^+, OH^-, K^+ as well as the components of phosphoric acid H_3PO_4, $H_2PO_4^-$, HPO_4^{2-}, and PO_4^{3-}.

As mentioned in Sec. II B, the proteins in this solution are assumed to have a fixed charge and can therefore be integrated with Eq. (5) with a chemical reaction source term $\hat{w}_i$ according to a model irreversible labeling reaction

$$\mathrm{U}+\mathrm{D}\xrightarrow{k_\mathrm{L}}\mathrm{L}. \tag{55}$$

The rate constant k_L in this reaction is pH dependent, given by the following:

$$k_\mathrm{L}=k_\mathrm{L}^0+d_\mathrm{L}e^{-(\mathrm{pH}-\mathrm{pH}_0)^2/\delta_\mathrm{pH}^2}. \tag{56}$$

The Gaussian dependence of this relationship on pH is based on the shape of the measured pH-dependence of the rate of production of the high-fluorescence-efficiency species from the reaction of naphthalene-2,3-dicarboxaldehyde (NDA) with amino acids in the presence of CN^-.[29] Unless stated otherwise, the values for the reaction rate parameters are

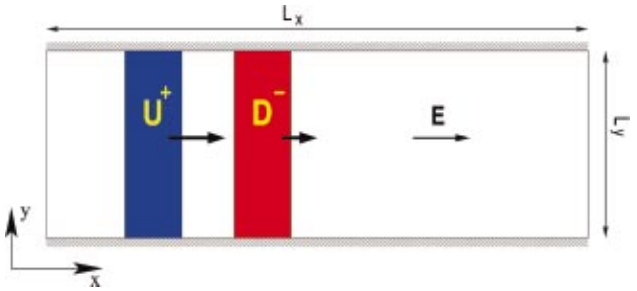

FIG. 2. (Color) Geometry for the numerical test problems: a plug of protein U and dye D are introduced in a rectangular microchannel and react to form a labeled protein L.

chosen in this work as $k_\mathrm{L}^0=0.25\times10^6$ mol^{-1} l s^{-1}, $d_\mathrm{L}=2.15\times10^6$ mol^{-1} l s^{-1}, $\mathrm{pH}_0=7.40$, and $\delta_\mathrm{pH}=0.85$. The chemical source terms used in Eq. (5) are correspondingly

$$\hat{w}_\mathrm{U}=\hat{w}_\mathrm{D}=-\hat{w}_\mathrm{L}=-k_L[\mathrm{U}][\mathrm{D}]. \tag{57}$$

The concentration of the K^+ ion, which is fully dissociated and is a conserved quantity can also be integrated by Eq. (5) directly (without a source term). Phosphoric acid, however, is a weak acid and will dissociate according to the following electrolyte reactions:

$$\mathrm{H_3PO_4}\xleftrightarrow{K_1}\mathrm{H^+}+\mathrm{H_2PO_4^-}, \tag{58}$$

$$\mathrm{H_2PO_4^-}\xleftrightarrow{K_2}\mathrm{H^+}+\mathrm{HPO_4^{2-}}, \tag{59}$$

$$\mathrm{HPO_4^{2-}}\xleftrightarrow{K_3}\mathrm{H^+}+\mathrm{PO_4^{3-}}, \tag{60}$$

where the K_i are the corresponding dissociation constants. As discussed in Sec. II B, an equilibrium formulation is used for these fast electrolyte reactions. Therefore, we consider the total concentration of this weak acid

$$\theta_a=[\mathrm{H_3PO_4}]+[\mathrm{H_2PO_4^-}]+[\mathrm{HPO_4^{2-}}]+[\mathrm{PO_4^{3-}}] \tag{61}$$

whose transport equation is obtained similarly to Eq. (10) by adding up the transport equations for all the components in θ_a so the dissociation reaction source terms disappear. The concentrations of the individual components of θ_a are then calculated as $c_i=\alpha_i\theta_a$, where the α_i are calculated from the equilibrium expressions for the dissociation reactions (58)–(60) and can be written as a function of $[\mathrm{H^+}]$ and the dissociation constants only:

$$\alpha_{\mathrm{H_3PO_4}}=\frac{[\mathrm{H^+}]^3}{[\mathrm{H^+}]^3+K_1[\mathrm{H^+}]^2+K_1K_2[\mathrm{H^+}]+K_1K_2K_3}, \tag{62}$$

$$\alpha_{\mathrm{H_2PO_4^-}}=\frac{\mathrm{K_1}[\mathrm{H^+}]^2}{[\mathrm{H^+}]^3+K_1[\mathrm{H^+}]^2+K_1K_2[\mathrm{H^+}]+K_1K_2K_3}, \tag{63}$$

$$\alpha_{\mathrm{HPO_4^{2-}}}=\frac{\mathrm{K_1K_2}[\mathrm{H^+}]}{[\mathrm{H^+}]^3+K_1[\mathrm{H^+}]^2+K_1K_2[\mathrm{H^+}]+K_1K_2K_3}, \tag{64}$$

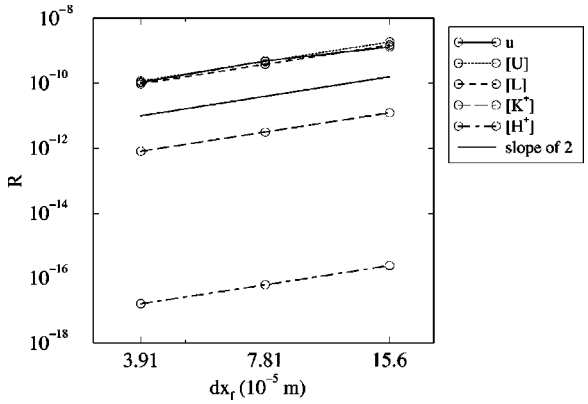

FIG. 3. L^2 norm of the difference between solutions on successive grids as a function of the fine grid spacing dx_f. The slope of the lines shows a second-order spatial convergence rate for various species concentrations as well as the streamwise velocity.

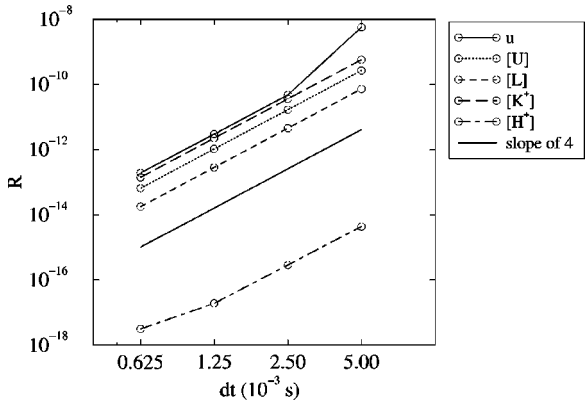

FIG. 4. L^2 norm of the difference between solutions at successive time steps as a function of the shorter time step dt. The slope of the lines shows a fourth-order temporal convergence rate for various species concentrations as well as the streamwise velocity.

$$\alpha_{PO_4^{3-}} = \frac{K_1K_2K_3}{[H^+]^3 + K_1[H^+]^2 + K_1K_2[H^+] + K_1K_2K_3}. \tag{65}$$

As discussed in Sec. II A, Eq. (4) is used to model the dependence of the zeta potential on pH and buffer molarity. The concentration of the fully dissociated potassium ion, $[K^+]$, is used for the local buffer molarity $\mathcal{M}$ along the walls. The temperature is assumed constant in this work, with all species properties and reaction rate constants evaluated at 298 K.

For the computations in this paper, all parameters and field variables, were represented with third-order polynomial chaos expansions. The highest order stochastic modes in the expansions of the predicted field variables were significantly lower than the lower order modes, indicating that the third-order expansions were sufficiently accurate.

B. Convergence with grid spacing

To test the spatial convergence rate of the code, simulations of the test case described above were run on a domain with $L_x = 1$ cm and $L_y = 0.25$ cm. The potassium phosphate buffer solution was initialized with a uniform concentration of 10^{-3} mol/l and a pH of 7.25. The unlabeled protein U and the dye D were initialized with a profile, Gaussian in x and uniform in y, both with a maximum concentration of 10^{-5} mol/l at $x = 4$ mm and a width of 1 mm. The labeled protein concentration was initialized to zero. The electrostatic potential difference ΔV between the inlet and exit of the domain was set to 10 V, creating an average field strength of 0.01 kV/cm. An uncertainty of 1% was assumed in the mobilities of both U and D, in the labeling rate parameter pH_0 of Eq. (56), and in the potential difference ΔV. Using third order polynomial chaos expansions, these four uncertain parameters led to four stochastic dimensions with a total of $P+1 = 35$ stochastic modes.

Four runs were performed, with uniform grid spacings in x and y doubling between each run, from 3.91×10^{-5} m in the finest grid to 3.13×10^{-4} m in the coarsest grid (corresponding, respectively, to 256×64, 128×32, 64×16, and 32×8 cells in $x\times y$). Each run used the same time step of 10^{-4} s for a total of 200 time steps. Figure 3 shows the L^2 norm of the difference between the solutions for the streamwise velocity u as well as several species concentrations at successive grid spacings. To monitor the spatial convergence of the full stochastic solution, the L^2 norm was calculated over all points in space and all $P+1$ stochastic modes. Clearly, the slope of the curves in Fig. 3 shows an overall second-order convergence rate with grid spacing, consistent with the spatial differencing scheme used.

C. Convergence with time step

The temporal convergence behavior of the code was studied with a similar test case as in the previous section. Referring to Fig. 2, the domain sizes were chosen as $L_x = 2$ cm and $L_y = 0.25$ cm. The buffer initialization was the same as in the previous case. For the unlabeled protein U and the dye D, however, the peak concentrations were raised to 10^{-4} mol/l, located at $x = 4$ mm and $x = 6$ mm, respectively. The electrostatic potential difference ΔV across the domain was set to 2000 V, giving an average field strength of 1 kV/cm. A slightly higher uncertainty of 2% was assumed in the mobilities of both U and D, the parameters pH_0 and ΔV, as well as the bulk kinematic viscosity. These five stochastic dimensions with third-order polynomial chaos expansions led to a total of $P+1 = 56$ stochastic modes.

This test case was run for a total time of 0.5 s, with five different time steps, ranging in factors of 2 from 6.25×10^{-4} s up to 1.00×10^{-2} s. In each case, the number of cells was 128×16 in $x\times y$. Figure 4 shows the L^2 norm of the difference between the solutions for the streamwise velocity u as well as several species concentrations at successive time steps. The fourth-order temporal convergence rate observed in Fig. 4 is consistent with the Runge–Kutta scheme used in the time integration.

VI. PROTEIN LABELING IN A HOMOGENEOUS BUFFER

To illustrate the stochastic uncertainty quantification methodology, this section describes protein labeling in a

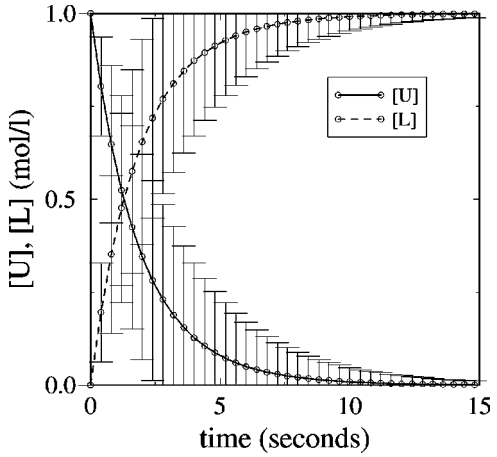

FIG. 5. Time evolution of U and L concentrations in a homogeneous protein labeling reaction. The uncertainty in these concentrations, due to a 1% uncertainty in the labeling reaction rate parameters, is indicated by $\pm 3\sigma$ ''error bars.''

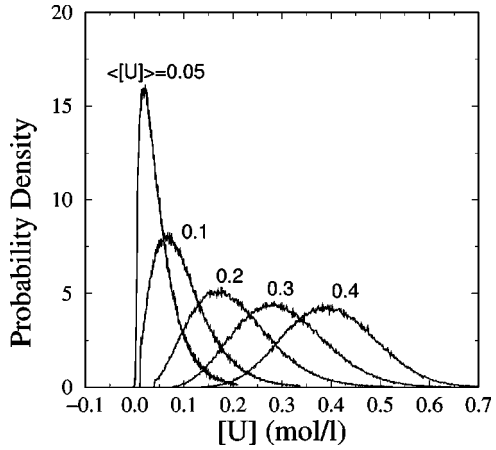

FIG. 6. PDF of the unlabeled protein concentration at different mean values. As the unlabeled protein reacts away, its PDF becomes narrower and more skewed.

simple homogeneous system. Figure 5 shows the time evolution of the concentrations of the unlabeled and labeled protein in a homogeneous potassium phosphate buffer at a pH of 8.25. In this problem, the dye D was assumed to be present in abundance so that the source term for the labeled protein in Eq. (5) can be written as

$$\hat{w}_{\mathrm{L}} = k_{\mathrm{L}}[\mathrm{U}]. \tag{66}$$

The same expression as before, Eq. (56), was used for the reaction rate, but with the following parameters: $k_{\mathrm{L}}^0 = 0.25 \times 10^{-3}\ \mathrm{s}^{-1}$, $d_{\mathrm{L}} = 2.15\ \mathrm{s}^{-1}$, $\mathrm{pH}_0 = 9.25$, and $\delta_{\mathrm{pH}} = 0.85$. Both proteins U and L, as well as the dye D were assumed to have no charge, and therefore the buffer equilibrium and pH did not change with time. For this simulation, a standard deviation of 1% was assumed for all parameters in the rate expression (56), as well as for the electrolyte dissociation constants. Third order PC expansions were used.

The resulting uncertainty in the protein concentrations is indicated in Fig. 5 with ''error bars'' that span the $\pm 3\sigma$ range, where σ indicates the standard deviation. Clearly, uncertainty in the input parameters causes large uncertainties in the simulated concentrations. At the point where $[\mathrm{U}] = 0.5$, a standard deviation of 1% in the parameter pH_0 is magnified about 16 times in the standard deviation of $[\mathrm{U}]$.

Note that after about 3 s, the range of the $\pm 3\sigma$ ''error bars'' becomes so large that it seems to include concentrations for U that are negative, which is clearly not physically possible. However, the interval $\pm 3\sigma$ around the mean value properly represents the full range of possibilities for a certain variable only when its probability density function is Gaussian, and therefore symmetric. Figure 6 shows the probability density function of $[\mathrm{U}]$, generated from its PC expansion at various points in time. When the mean value of $[\mathrm{U}]$ is sufficiently far away from zero, this PDF has a Gaussian shape. However, for mean values of $[\mathrm{U}]$ closer to zero, the PDF becomes narrower and more skewed. This predicted uncertainty properly reflects the physical system behavior where all unlabeled protein reacts away, but its concentration can not be negative.

VII. PROTEIN LABELING IN A TWO-DIMENSIONAL MICROCHANNEL

In this section, the simulation and uncertainty quantification code is used to tackle a more physically challenging problem of protein labeling in a two-dimensional microchannel. The problem setup is similar to the numerical test problems described in Sec. V. The labeling reaction is the same as Eq. (55) with the reaction rate k_L and the corresponding source terms as in Eqs. (56) and (57). Again, a charge of $+1$ is assumed for the unlabeled protein U and a charge of -1 for the dye D, resulting in a neutral labeled protein L.

Referring to Fig. 2, a microchannel was considered with a length $L_x = 1$ cm and a height $L_y = 1$ mm. The potassium phosphate buffer solution was initialized with a uniform concentration of 10^{-3} mol/l and a pH of 7.25. The Gaussian profiles for the initial U and D concentrations had peak concentrations of 10^{-4} mol/l, located at $x = 2.5$ mm and $x = 4$ mm, respectively, and a width in x of 0.75 mm. The electrostatic potential difference ΔV across the domain was set to 1000 V, giving an average field strength of 1 kV/cm. An uncertainty of 1% was assumed in the mobility of U, in the labeling rate parameter pH_0, the dissociation constant K_2, and the potential difference ΔV. Third-order polynomial chaos expansions were used in the computations with a total of 35 stochastic modes. The time step was set to 2×10^{-4} s and the domain was discretized with 512×32 cells in x and y.

Figure 7 shows a contour plot of the mean concentrations of the proteins and dye at $t = 0.12$ s. At this point in time, the plugs of U and D have just met at $x \approx 4$ mm, and labeled protein is generated at the interface. Note that the labeling reaction is fast compared to the electroosmotic and electrophoretic transport. Consequently, U and D react as soon as they meet, resulting in almost no overlap between the U and D profiles, and a sharp profile for L. Since L is neutral, it travels with the bulk convective velocity, which is the average of the total convective velocities of U and D. Therefore the peak value of L is always located at the interface of U and D, and since L is generated in that same area,

its peak concentration will keep increasing. At $t=0.12$ *s*, the peak concentration for L is 1.3×10^{-4} mol/l, which is already higher than the peak concentrations of 9.4×10^{-5} mol/l for U and D.

The standard deviations in the concentrations of Fig. 7 are given in Fig. 8. The highest uncertainties appear in the reaction zone at the interface between U and D, with a maximum coefficient of variation of about 20% in the L concentration. Even though Fig. 8 only shows the overall uncertainty in the concentrations, a strong feature of the PC formalism is that the contributions of the uncertainty in individual parameters to this overall uncertainty can easily be retrieved, as explained in Sec. III. Figure 9, for example, shows the contributions from each of the four uncertain input parameters to the standard deviation of the L concentration, in the area around the reaction zone, at $y=0.5$ mm. These contributions were obtained with a similar analysis as in Eqs. (27)–(32), but for the case of four stochastic dimensions and third-order PC expansions. The total standard deviation of [L] is given by the curve labeled ''all'' in this figure. This overall standard deviation has a profile with a double peak, which for a single peak mean species profile, is characteristic of uncertainty caused by the convection velocity. When a single peak species profile is transported by an uncertain convection velocity, the uncertainty in the position of the peak at a given point in time will cause the most variability at the sides of the peak, where the profile has a steep slope in the *x* direction. At the top of the profile, there is no concentration gradient and uncertainties in peak position cause little uncertainty in the observed concentrations at that location.

As indicated by the curve labeled ''ΔV,'' the uncertainty in the applied electrostatic field potential has the most dominant contribution to the overall standard deviation. Since both the electroosmotic and electrophoretic velocities are directly proportional to ΔV, the uncertainty caused by this parameter naturally shows a double peak, characteristic of convection velocity uncertainty. Similarly, the parameter β_{U} affects the electrophoretic transport of the reactant U and its resulting contribution to the standard deviation of [L] also has a double peak, albeit smaller than the ΔV contribution.

The contribution of parameter pH_0 also shows a double peak, but with its center located on the left side of the [L] profile, where the gradient of [L] in *x* is very steep. The steepness of the [L] profile in that area is largely determined by the speed of the labeling reaction compared to the convection speed, with a faster reaction rate leading to a sharper increase in [L]. With the pH in this area between 7.0 and 7.1 (not shown), Eq. (56) predicts significant variability in k_{L} for changes in pH_0. So the uncertainty in pH_0 mainly affects the slope of the [L] profile on the left side, consistent with the observed contribution of parameter pH_0 in Fig. 9.

Figure 9 further shows more minor contributions, from the dissociation parameter K_2 and from the coupled terms. Even though their contribution is small in this case, those coupled terms are interesting from a theoretical point of view, as they represent coupled effects of independent parameters. In the current figure, those terms represent the sum of three different coupled effects: the coupled effect of ΔV and β_U, of ΔV and pH_0, and of ΔV and K_2.

As time goes on and the U and D plugs cross each other, nearly all U and D are consumed in the labeling reaction. At $t=0.50$ s, only labeled protein L remains, with its mean concentration and standard deviation as shown in Fig. 10. The maximum mean concentration of L at this point in time is 2.4×10^{-4} mol/l in the center of the channel, and about 3.2×10^{-4} mol/l near the walls. So the L concentration is up to three times as large as the initial U and D concentrations. The standard deviation in L, as shown in the bottom plot of Fig. 10, is very large near the wall, with maximum values up to 10^{-4} mol/l and coefficients of variation up to 100%. Again, the standard deviation in [L] exhibits the double peak near the centerline, which is characteristic of uncertainty caused by the convection velocity.

What is particularly significant though, is the major distortion of the L plug, as opposed to the straight profile observed at early times. This distortion is caused by the disturbance of the buffer electrolyte, in response to the movement and annihilation of the charged protein U and the dye D. To explain why this is physically happening, consider Fig. 11, which shows the mean and standard deviation of the electrical conductivity σ of the electrolyte solution at $t=0.50$ s. Because two charged molecules are used up for every new labeled protein, the area around the L plug has a reduced concentration of ions, with a mean electrical conductivity of almost a third lower than in the undisturbed buffer. Upstream of the L plug, the electrical conductivity shows some smaller fluctuations, which stem from shifts in the buffer equilibrium. Since the buffer ions are primarily negatively charged, those disturbances travel slower than the labeled protein plug. The bottom plot of Fig. 11 shows that the highest uncertainties in the electrical conductivity are found around the L plug, near the center and especially at the walls.

The large spatial variations in the electrical conductivity in turn cause nonuniformities in the electrical field strength, as shown in Figs. 12 and 13. Near the L plug, the mean electrostatic field strength in the *x* direction reaches a value up to 40% higher than in the undisturbed flow. This increase strongly affects the local electroosmotic and electrophoretic velocities, causing an increased wall velocity, leading to the observed distortion of the L plug. The largest uncertainties are again found near the L plug, with maxima up to 10%. Even though the initial field strength in the *y* direction was zero, Fig. 13 shows that this *y* component is quite significant at $t=0.50$ s. The magnitude of this field strength is up to 15% of the initial, streamwise electrostatic field strength for the mean value. Even though this *y* component does not affect the electroosmotic flow velocity directly, it does provide electrophoretic ion transport in the wall-normal direction, which can further distort sample profiles.

As indicated by Eq. (3), the electroosmotic wall velocity depends on both the local electrostatic field strength and ζ potential, which in turn depends on the pH and the buffer molarity, as modeled by Eq. (4). Since all these variables are disturbed by the charged protein movement and annihilation, the electroosmotic wall velocity varies in the streamwise direction. These wall velocity changes in turn cause pressure gradients and local recirculation zones, as indicated by the velocity fields in Figs. 14 and 15. Figure 14 shows the

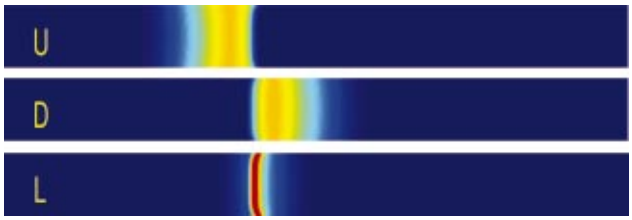

FIG. 7. (Color) Mean concentrations of proteins U, L, and dye D at $t=0.12$ s. U and D just met and L is produced at their interface. The values of the contour levels go linearly from 0 (blue) to 1.3×10^{-4} mol/l (red). In this figure, as well as in all subsequent contour plots, the full physical domain is shown, from 0 to 1 cm in x and from 0 to 1 mm in y.

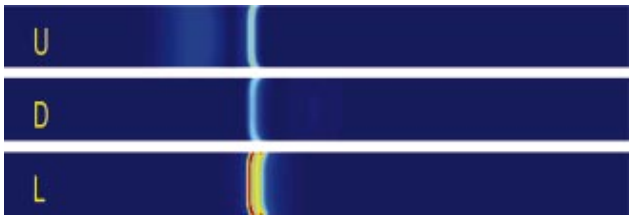

FIG. 8. (Color) Standard deviation of the protein and dye concentrations at $t=0.12$ s. The values of the contour levels go linearly from 0 (blue) to 1.1×10^{-5} mol/l (red). The largest uncertainties are found in the reaction zone.

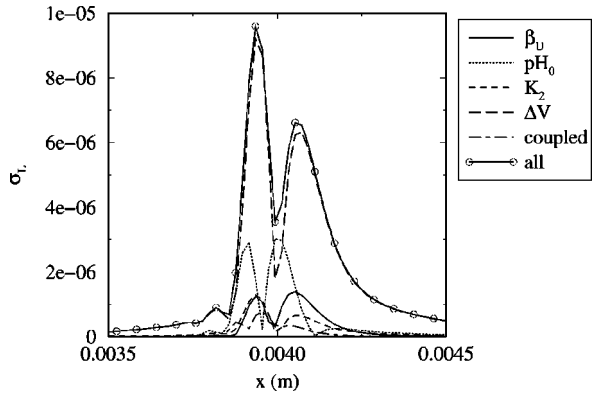

FIG. 9. Major contributions of individual input parameters to the overall standard deviation in [L] in the area around the reaction zone at $t=0.12$ s, $y=0.5$ mm. The uncertainty in the applied voltage potential "ΔV" has the most dominant contribution to the overall standard deviation in [L].

FIG. 10. (Color) Mean (top) and standard deviation (bottom) of the labeled protein concentration L at $t=0.50$ s. The initially flat profiles are now severely distorted. The values of the contour levels go linearly from 0 (blue) to 3.2×10^{-4} mol/l (red) in the top plot and from 0 (blue) to 10^{-4} mol/l (red) in the bottom plot.

FIG. 11. (Color) Mean (top) and standard deviation (bottom) of the electrical conductivity of the electrolyte solution at $t=0.50$ s. Annihilation of ions in the labeling reaction results in a significantly lower mean electrical conductivity near the L plug. The values of the contour levels go linearly from 7.1×10^{-3} S/m (blue) to 1.3×10^{-2} S/m (red) in the top plot and from 0 (blue) to 1.5×10^{-3} S/m (red) in the bottom plot.

FIG. 12. (Color) Mean (top) and standard deviation (bottom) of the electrical field strength in the x direction at $t=0.50$ s. Near the L plug, the mean streamwise electrical field strength is about 40% higher than in the undisturbed flow. The values of the contour levels go linearly from 91.4 kV/m (blue) to 146 kV/m (red) in the top plot and from 0.20 kV/m (blue) to 13 kV/m (red) in the bottom plot.

FIG. 13. (Color) Mean (top) and standard deviation (bottom) of the electrical field strength in the y direction at $t=0.50$ s. The magnitude of the mean of this field strength is up to 15% of the initial field strength in the x direction. The values of the contour levels go linearly from -16.3 kV/m (blue) to 16.3 kV/m (red) in the top plot and from 0 (blue) to 5.8 kV/m (red) in the bottom plot.

FIG. 14. (Color) Mean (top) and standard deviation (bottom) of the streamwise velocity at $t=0.50$ s. The local increase in the electroosmotic wall velocity leads to recirculation zones near the L plug. The largest uncertainties are found near the wall. The values of the contour levels go linearly from 6.8 mm/s (blue) to 9.1 mm/s (red) in the top plot and from 5.6×10^{-3} mm/s (blue) to 0.59 mm/s (red) in the bottom plot.

FIG. 15. (Color) Mean (top) and standard deviation (bottom) of the wall-normal velocity at $t=0.50$ s. The mean of this velocity has a magnitude of up to 6 % of the initial streamwise velocity. The values of the contour levels go linearly from -0.56 mm/s (blue) to 0.56 mm/s (red) in the top plot and from 0 (blue) to 0.26 mm/s (red) in the bottom plot.

streamwise velocity field, which has a mean wall velocity that is up to 20% higher near the L plug. The wall-normal velocity field shows positive and negative velocities near the L plug, with magnitudes up to 6% of the initial streamwise velocity.

Clearly, the recirculation zones in the flow field will distort initially flat sample profiles. This increases the hydrodynamic dispersion, on top of the electrokinetic dispersion caused by nonuniformities in the electrophoretic transport.

VIII. CONCLUSIONS

In this paper, a detailed physical model for microchannel flows was presented to study protein labeling reactions in an electrolyte buffer. The model incorporates the coupled nature of momentum transport, species transport, and the electrostatic field as well as a full representation of the electrolyte buffer reactions and the dependence of the ζ potential on the local buffer properties. A stochastic uncertainty quantification method was developed to propagate uncertainty from the input parameters in the model to the simulation results, using polynomial chaos expansions for the uncertain model parameters and field variables.

Application of the model was illustrated in light of simulations of protein labeling reactions in homogeneous systems as well as two-dimensional electroosmotically driven microchannel flows. For the two-dimensional case, the simulation showed the impact of ion movement and subsequent buffer disturbances on the electrokinetic and hydrodynamic dispersion of sample plugs in the channel. The uncertainty in the results was primarily due to uncertainty in the applied voltage across the channel, with smaller contributions from the parameters in the labeling reaction rate as well as species properties.

Overall, the detailed physical model that was implemented in this work, allows the simulations of microchannel flows providing in-depth understanding of the transport and dispersion of protein sample plugs. In combination with this model, the stochastic uncertainty quantification method provides a powerful way to assess the impact of uncertain model input parameters on the uncertainty of the simulation results.

ACKNOWLEDGMENTS

This work was supported by the Defense Advanced Research Projects Agency (DARPA) and Air Force Research Laboratory, Air Force Materiel Command, USAF, under Agreement No. F30602-00-2-0612. The U.S. government is authorized to reproduce and distribute reprints for governmental purposes notwithstanding any copyright annotation thereon.

[1]W. Thormann and R. Mosher, ''Theory of elecrophoretic transport and separations: The study of electrophoretic boundaries and fundamental separation principles by computer simulation,'' in *Advances of Electrophoresis*, edited by A. Chrambach, M. J. Dunn, and B. J. Radola (VCH, Weinheim, Germany, 1998), Vol. 2, pp. 45–108.

[2]G. Roberts, P. Rhodes, and R. Snyder, ''Dispersion effects in capillary zone electrophoresis,'' J. Chromatogr. **480**, 35 (1989).

[3]M. Bier, O. Palusinski, R. Mosher, and D. Saville, ''Electrophoresis: Mathematical modeling and computer simulation,'' Science **219**, 1281 (1983).

[4]O. Palusinski, A. Graham, R. Mosher, M. Bier, and D. Saville, ''Theory of elecrophoretic separation. Part II: Construction of a numerical simulation scheme and its applications,'' AIChE J. **32**, 215 (1986).

[5]S. Ermakov and P. Righetti, ''Computer simulation for capillary zone electrophoresis, a quantitative approach,'' J. Chromatogr., A **667**, 257 (1994).

[6]C. Shafer-Nielsen, ''A computer model for time-based simulation of electrophoresis systems with freely defined initial and boundary conditions,'' Electrophoresis **16**, 1369 (1995).

[7]V. McGuffin and M. Tavares, ''Computer assisted optimization of separations in capillary zone electrophoresis,'' Anal. Chem. **69**, 152 (1997).

[8]W. Thormann, C.-X. Zhang, J. Caslavska, P. Gebauer, and R. Mosher, ''Modeling of the impact of ionic strength on the electroosmotic flow in capillary electrophoresis with uniform and discontinuous buffer systems,'' Anal. Chem. **70**, 549 (1998).

[9]N. Patankar and H. Hu, ''Numerical simulation of electroosmotic flow,'' Anal. Chem. **70**, 1870 (1998).

[10]S. Ermakov, S. Jacobson, and J. Ramsey, ''Computer simulations of electrokinetic transport in microfabricated channel structures,'' Anal. Chem. **70**, 4494 (1998).

[11]S. Ermakov, S. Jacobson, and J. Ramsey, ''Computer simulations of electrokinetic injection techniques in microfluidic devices,'' Anal. Chem. **72**, 3512 (2000).

[12]Q. Mao, J. Pawliszyn, and W. Thormann, ''Dynamics of capillary isoelectric focusing in the absence of fluid flow: High-resolution computer simulation and experiment validation with whole column optical imaging,'' Anal. Chem. **72**, 5493 (2000).

[13]V. Andreev, N. Pliss, and P. Righetti, ''Computer simulation of affinity capillary electrophoresis,'' Electrophoresis **23**, 889 (2002).

[14]R. Ghanem and P. Spanos, *Stochastic Finite Elements: A Spectral Approach* (Springer, New York, 1991).

[15]N. Wiener, ''The homogeneous chaos,'' Am. J. Math. **60**, 897 (1938).

[16]O. Le Maître, M. Reagan, H. Najm, R. Ghanem, and O. Knio, ''A stochastic projection method for fluid flow II. Random process,'' J. Comput. Phys. **181**, 9 (2002).

[17]R. Probstein, *Physicochemical Hydrodynamics: An Introduction*, 2nd ed (Wiley, New York, 1994).

[18]Q.-H. Wan, ''Effect of electroosmotic flow on the electrical conductivity of packed capillary columns,'' J. Phys. Chem. B **101**, 4860 (1997).

[19]P. Scales, F. Grieser, T. Healy, L. White, and D. Chan, ''Electrokinetics of the silica-solution interface: A flat plate streaming potential study,'' Langmuir **8**, 965 (1992).

[20]R. Mosher, D. Dewey, W. Thormann, D. Saville, and M. Bier, ''Computer simulation and experiment validation of the electrophoretic behavior of proteins,'' Anal. Chem. **61**, 362 (1989).

[21]R. Mosher, P. Gebauer, and W. Thormann, ''Computer simulation and experimental validation of the electrophoretic behavior of proteins. III. Use of titration data predicted by the protein's amino acid composition,'' J. Chromatogr. **638**, 155 (1993).

[22]S. Ermakov, M. Zhukov, L. Capelli, and P. Righetti, ''Artifactual peak splitting in capillary electrophoresis. 2. Defocusing phenomena for ampholytes,'' Anal. Chem. **67**, 2957 (1995).

[23]O. Le Maître, O. Knio, R. Ghanem, and H. Najm, ''A stochastic projection method for microchannel flow,'' in Proceedings of the Fourth International Conference on Modeling, 2001, pp. 246–249.

[24]O. Le Maître, O. Knio, H. Najm, and R. Ghanem, ''A stochastic projection method for fluid flow I. Basic formulation,'' J. Comput. Phys. **1732**, 481 (2001).

[25]M. Pernice and H. Walker, ''NITSOL: A Newton iterative solver for nonlinear systems,'' SIAM J. Sci. Comput. (USA) **19**, 302 (1998).

[26]A. Chorin, ''A numerical method for solving incompressible viscous flow problems,'' J. Comput. Phys. **2**, 12 (1967).

[27]E. Hairer, S. Nørsett, and G. Wanner, *Solving Ordinary Differential Equations I, Nonstiff Problems*, 2nd ed. (Springer, Berlin, 1992).

[28]SLATEC common mathematical library, version 4.1, http://www.netlib.org/slatec/, 1993.

[29]P. de Montigny, J. Stobaugh, R. Givens, R. Carlson, K. Srinivasachar, L. Sternson, and T. Higuchi, ''Naphtalene-2,3-dicarboxaldehyde/cyanide ion: A rationally design fluorogenic reagent for primary amines,'' Anal. Chem. **59**, 1096 (1987).